



TRADUÇÃO

*Machine learning e o direito*¹⁻²

Machine learning and law

Machine learning y derecho

Harry Surden³

University of Colorado Law School (Colorado, EUA)

ORCID: <https://orcid.org/0000-0003-4434-1926>

E-mail: hsurden@colorado.edu

Saul Tourinho Leal⁴

Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (Brasília, DF, Brasil)

ORCID: <https://orcid.org/0000-0002-8816-4514>

E-mail: stourinho@ayresbritto.com.br

Wilson Seraine da Silva Neto⁵

Faculdade de Direito da Universidade de Coimbra (Coimbra, Portugal)

ORCID: <https://orcid.org/0000-0001-8227-8265>

E-mail: serainenetow@gmail.com

¹ SURDEN, Harry. Machine learning e o direito. Tradução de Saul Tourinho Leal, Wilson Seraine da Silva Neto. **Suprema**: revista de estudos constitucionais, Brasília, v. 3, n. 1, p. 353-389, jan./jun. 2023. DOI: <https://doi.org/10.53798/suprema.2023.v3.n1.a212>.

² A presente tradução foi realizada de forma independente pelos tradutores, com a devida autorização do autor, Harry Surden, e da *Editor-in-Chief da Washington Law Review*, Hannah Garland, sem, contudo, participação, seja de revisão ou de supervisão, desses.

³ Professor na *University of Colorado Law School*. Foi *law clerk* do juiz Martin J. Jenkins no *United States District Court for the Northern District of California*, em San Francisco. É o diretor associado do *Stanford Center for Legal Informatics (CodeX)*. <https://www.harrysurden.com/wordpress/about>

⁴ Pós-doutor em Direito Constitucional pela Universidade Humboldt, sob a orientação de Dieter Grimm. Foi assessor estrangeiro na Corte Constitucional da África do Sul e da vice-presidência da Suprema Corte de Israel. Advogado. Currículo Lattes: <http://lattes.cnpq.br/3932505807469261>

⁵ Mestrando em Ciências Jurídico-Políticas, com menção em Direito Constitucional, pela Faculdade de Direito da Universidade de Coimbra. Currículo Lattes: Advogado. <http://lattes.cnpq.br/4454693212968845>

Sumário

1. Introdução. 2. Uma breve síntese do *machine learning*. 2.1 O que é *machine learning*? 2.1.1 Filtros de *e-mails spam* como exemplo de *machine learning*. 2.1.2 Detecção de padrões para modelar fenômenos complexos. 3. Resultados inteligentes sem a inteligência. 3.1 *Proxies* e heurísticas para inteligência. 3.1.1 Assemelhando-se à inteligência por *proxy*. 3.1.2 Os desenvolvimentos nas pesquisas de Inteligência Artificial. 4. *Machine learning* e direito. 4.1 *Machine Learning* aplicado ao direito. 4.2 Modelos preditivos. 4.2.1 Previsões judiciais. 4.2.2 Limitações aos modelos preditivos judiciais de *machine learning*. 4.3 Relações ocultas nos dados. 4.3.1 Decisões judiciais e relacionamentos de dados. 4.4 Classificação e agrupamento de documentos. 4.4.1 Classificação automatizada de documentos. 4.4.2 Classificação de documentos de processos contenciosos. 4.4.3 *E-discovery* e classificação de documentos. 4.4.4 Agrupamento de documentos relacionados. 5. Conclusão

Contents

1. Introduction. 2. Overview of machine learning. 2.1 What is machine learning? 2.1.1 Email spam filters as an example of machine learning. 2.1.2 Detecting patterns to model complex phenomena. 3. Intelligent results without intelligence. 3.1 Proxies and heuristics for intelligence. 3.1.1 Approximating intelligence by proxy. 3.1.2 Developments in AI research. 4. Machine learning and law. 4.1 Machine learning applied to law. 4.2 Predictive Models. 4.2.1 Legal Predictions. 4.2.2 Limitations to machine learning legal predictive models. 4.3 Finding hidden relationships in data. 4.3.1 Judicial decisions and data relationship. 4.4 Document Classification and Clustering. 4.4.1 Automated document classification. 4.4.2 Classification of litigation docket documents. 4.4.3 E-discovery and document classification. 4.4.4 Clustering and grouping of related documents. 5. Conclusion.

Índice

1. Introducción. 2. Una breve descripción del *machine learning*. 2.1 ¿Qué es el *machine learning*? 2.1.1 Filtros de *e-mails spam* como ejemplo de *machine learning*. 2.1.2 Detección de patrones para modelar fenómenos complejos. 3. Resultados inteligentes sin inteligencia. 3.1 *Proxies* y heurísticas para inteligencia. 3.1.1 Aproximación de la inteligencia por *Proxy*. 3.1.2 Desarrollos en la investigación de Inteligencia Artificial. 4. *Machine learning* y derecho. 4.1 *Machine learning* aplicado al derecho. 4.2 Modelos predictivos. 4.2.1 Pronóstico judiciales. 4.2.2 Limitaciones de los modelos predictivos judiciales de *machine learning*. 4.3 Relaciones ocultas en los datos. 4.3.1 Decisiones judiciales y relaciones de datos. 4.4. Clasificación y agrupación de documentos. 4.4.1 Clasificación automatizada de documentos. 4.4.2 Clasificación de

los documentos del proceso litigioso. 4.4.3 *E-discovery* y clasificación de documentos. 4.4.4 Agrupación de documentos relacionados. 5. Conclusión

1. Introdução

Qual o impacto da Inteligência Artificial (IA) sobre a prática do direito? Se atermos a um determinado entendimento, a IA deveria ter pouca influência sobre a *práxis* forense, exceto com avanços técnicos significativos⁶, uma vez que essa requer habilidades cognitivas avançadas, ainda fora da capacidade do atual estágio da IA⁷. Advogados, por exemplo, combinam cotidianamente raciocínio abstrato e habilidades de resolução de problemas em ambientes de incerteza, tanto jurídica quanto factual⁸. Os algoritmos de IA modernos, ao contrário, não conseguem, no estágio atual, replicar a maioria das habilidades intelectuais humanas, ficando muito aquém dos processos cognitivos avançados, como o raciocínio analógico, que é indispensável para a prática jurídica⁹. Diante dessas e outras limitações na atual tecnologia de IA, pode-se concluir que, até que os computadores possam replicar o avançado processo cognitivo rotineiramente exercido por advogados experientes, a IA teria pouco¹⁰ impacto em um domínio tão cheio de abstração e incerteza como o direito.

Contudo, embora haja alguma verdade nesse entendimento, sua conclusão é excessivamente genérica, pois desconsidera uma variedade de tarefas jurídicas para as quais a atual tecnologia de IA pode ter um impacto, mesmo com a sua ainda incapacidade tecnológica, se compararmos ao raciocínio humano. Considere que, fora do direito, técnicas de IA não cognitivas foram aplicadas com sucesso em atividades que antes se pensava ser necessária a inteligência humana – como por

⁶ Ver, e.g., “Symposium, Legal Reasoning and Artificial Intelligence: How Computers “Think” Like Lawyers”, 8 U. Chi. L. Sch. Roundtable 1, 19 (2001) (Cass Sunstein afirma que: “No atual estado da arte, a Inteligência Artificial não consegue se empenhar em um raciocínio analógico ou raciocínio jurídico”).

⁷ Ver, e.g., Karl Okamoto, Teaching Transactional Lawyering, 1 Drexel L. Rev. 69, 83 (2009) (“A essência da advocacia é a “resolução criativa de problemas” sob condições de incerteza e complexidade. Essa concepção de advocacia como solução de problemas tornou-se senso comum.”).

⁸ Karl Okamoto, Teaching Transactional Lawyering, p. 83

⁹ Karl Okamoto, Teaching Transactional Lawyering, p. 83

¹⁰ Ver: Harry Surden, Computable Contracts, 46 U.C. Davis L. Rev. 629, 646 (2012) (discute-se como as mudanças de linguagem, que são normalmente triviais para os humanos entenderem, podem confundir os algoritmos do computador).

exemplo, tradução de idiomas¹¹. Embora os resultados desses esforços automatizados às vezes sejam imperfeitos, o ponto interessante é que esses resultados gerados por computador muitas vezes se mostram úteis para tarefas específicas, em que um certo grau de aproximação é aceitável.¹² Nessa mesma linha, este artigo irá sugerir que há um subconjunto limitado, mas não insignificante, de tarefas jurídicas que podem ser parcialmente automatizadas, a partir das técnicas atuais de IA, apesar de suas limitações quando comparada à cognição humana.

Em particular, este artigo dedica-se a uma classe específica de IA, conhecida como *machine learning* [em português, aprendizado (ou aprendizagem) de máquina¹³], e o seu potencial impacto na prática jurídica. Em poucas palavras, o *machine learning* consiste em algoritmos de computador que têm a capacidade de “aprender” ou aperfeiçoar, ao longo do tempo, o seu desempenho em alguma atividade.¹⁴ Sabendo-se das várias vertentes da IA, por que destacar o *machine learning* em particular? Nas últimas décadas, os pesquisadores usaram, com sucesso, o *machine learning* para automatizar uma variedade de tarefas sofisticadas que anteriormente se presumia exigir a cognição humana. Essas aplicações variam desde carros autônomos até a tradução de idiomas, previsão, reconhecimento da fala e visão computadorizado.¹⁵ Não diferentemente, os pesquisadores também começaram a aplicar essas técnicas no contexto do direito.¹⁶

Para ser claro, não estou sugerindo que todas as tarefas, ou mesmo a maioria, executadas rotineiramente pelos advogados sejam automatizáveis, haja vista o estado atual da tecnologia de IA, pois muitos dos trabalhos executados pelos advogados exigem o tipo de habilidades intelectuais que estão além da capacidade das tecnologias vigentes. Em vez disso, estou sugerindo que existem subconjuntos de atividades jurídicas que, provavelmente, são automatizáveis, sob o atual estado da

¹¹ Ver: David Bellos, “Is that a fish in your ear? Translation and the meaning of everything”, 253–57 (2011); Find Out How Our Translations Are Created, GOOGLE, <http://translate.google.com/about> (last visited Feb. 24, 2014).

¹² David Bellos, *Is That A Fish In Your Ear?*...

¹³ N.T: Em razão da notoriedade e da familiaridade do termo estrangeiro na literatura brasileira, optou-se por utilizar a nomenclatura original.

¹⁴ Peter Flach, “Machine learning: the art and science of algorithms that make sense of data 3” (2012).

¹⁵ Burkhard Bilger, “Auto Correct: Has the Self-Driving Car at Last Arrived?”, *NEW YORKER*, Nov. 25, 2013, em 96, 106; Parag Kulkarni, “Reinforcement and systemic machine learning for decision making”, p. 1–2 (2012) (discute-se a visão computadorizada).

¹⁶ Ver, e.g., Daniel Martin Katz, “Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry”, 62 *Emory L.J.* 909, 936 (2013) (discutem-se aplicações no direito, como automação na descoberta de documentos e previsão legal quantitativa).

arte, de modo a auxiliar nos afazeres mais relevantes, com a devida compreensão e contabilização das limitações de sua precisão. Em outras palavras, mesmo com as limitações atuais da IA, em comparação com a cognição humana, consegue-se produzir resultados que são “bons o suficiente” para determinados contextos jurídicos.

Nessa toada, a primeira parte deste artigo explicará os conceitos básicos subjacentes ao *machine learning*. Em seguida, transmitirá um princípio mais geral: algoritmos de computador não inteligentes podem, às vezes, produzir resultados inteligentes em tarefas complexas por meio do uso de determinados *proxies* detectados nos dados. E, por fim, demonstrar-se-á como certas tarefas jurídicas podem ser passíveis de automação parcial, sob esse princípio, através de *machine learning*. Ademais, também enfatizará as limitações significativas desses métodos automatizados em comparação com as capacidades de advogados em situações semelhantes.

2. Uma breve síntese do *machine learning*

2.1 O que é *machine learning*?

“*Machine Learning*” refere-se a um subcampo da ciência da computação que concerne em programas de computador capazes de aprender com a experiência e, assim, melhorar seu desempenho ao longo do tempo¹⁷. Como será discutido, a ideia de que os computadores estão “aprendendo” é, em grande parte, uma metáfora e não implica que os sistemas de computadores estejam replicando artificialmente os sistemas cognitivos avançados que há no aprendizado humano.¹⁸ Na verdade, consideramos esses algoritmos como sendo aprendizes no sentido *funcional*: eles são capazes, a partir da mudança do seu comportamento, de melhorar o desempenho na realização de alguma tarefa por meio da experiência.¹⁹

Geralmente, os algoritmos do *machine learning* são usados para detectar padrões em dados, com o escopo de automatizar tarefas complexas ou fazer previsões.²⁰ Atualmente, esses algoritmos são usados em uma variedade de aplicativos

¹⁷ Stuart Russell & Peter Norvig, “Artificial intelligence: a modern approach”, 693 (3d ed. 2010).

¹⁸ I. H. Witten, “Data mining: practical machine learning tools and techniques”, § 1.3 (3d ed. 2011).

¹⁹ I. H. Witten, “Data mining...”, § 1.3

²⁰ David E. Sorkin, “Technical and Legal Approaches to Unsolicited Electronic Mail”, 35 U.S.F. L. Rev. 325, 326 (2001).

comerciais do mundo real, incluindo resultados de pesquisa na Internet, reconhecimento facial, detecção de fraude e mineração de dados.²¹ O *machine learning* está mais intimamente associado à sua aplicação em “análises preditivas”, pois os pesquisadores geralmente empregam esses métodos algorítmicos para analisar dados existentes com o objetivo de prever a probabilidade de resultados futuros incertos.²² Se obtiverem um bom desempenho, os algoritmos de *machine learning* podem produzir resultados automatizados que se aproximam daqueles que teriam sido alcançados por uma pessoa em situação semelhante. Logo, esse subcampo da computação é, portanto, frequentemente considerado um ramo da inteligência artificial, uma vez que um algoritmo de bom desempenho pode produzir resultados automatizados que parecem “inteligentes”.²³

Assim, o objetivo desta primeira parte é demonstrar alguns princípios básicos do *machine learning* de maneira acessível ao público não técnico, a fim de expressar um ponto mais amplo sobre a possível aplicabilidade dessas técnicas a tarefas dentro do direito.

2.1.1 Filtros de *e-mails spam* como exemplo de *machine learning*

Para uma melhor explicação, utilizemos, como um exemplo familiar que ilustrará alguns recursos básicos comuns às técnicas de *machine learning*, os filtros de “*spam*” dos *e-mails*. O “*spam*” consiste em *e-mails* comerciais não solicitados e indesejados, que podem interferir no acesso dos usuários a mensagens tidas como mais importantes.²⁴ Em tese, o próprio usuário poderia gerenciar o *spam* manualmente ao ler cada *e-mail* e, ao identificar se uma determinada mensagem é ou não *spam*, excluí-la. No entanto, dado que essa tarefa é trabalhosa, seria desejável automatizar a identificação de *spam*. Assim, para executar essa filtragem automatizada, os programas de *software* de *e-mail* frequentemente usam algoritmos de *machine learning*.²⁵

Contudo, indaga-se: como os algoritmos de *machine learning* identificam automaticamente um *spam*? Esses algoritmos são projetados para detectar padrões entre os dados. Em um processo típico, esses algoritmos são “treinados” para

²¹ H. Witten, *Data Mining*, p. 12, no § 1.3.

²² Ver, e.g., Lawrence Maisel, “Predictive business analytics: forward looking capabilities to improve business performance”, 27–30 (2014).

²³ Stuart Russell & Peter Norvig, *Artificial Intelligence*, p. 3–5.

²⁴ David E. Sorkin, *Technical and Legal Approaches to Unsolicited Electronic Mail*, p. 325–30.

²⁵ David E. Sorkin, *Technical and Legal Approaches to Unsolicited Electronic Mail*, p. 325–30

reconhecer tais *e-mails* através de exemplos, possibilitando-os realizar uma análise de padrões. Isso ocorre, por exemplo, quando uma pessoa identifica um determinado *e-mail* como *spam* ao sinalizá-lo através do seu *software* de leitura de *e-mail*. Dessa forma, podemos pensar nesse ato de sinalizar como uma indicação para o algoritmo do computador de que esse é um exemplo verificado de que um *e-mail* é *spam*, devendo-se avaliar esse determinado padrão.²⁶

Dessa maneira, ao analisar o *e-mail* de *spam*, o algoritmo de *machine learning* irá tentar detectar as características que indicam que um determinado *e-mail* tem mais probabilidade de ser um *spam*. Depois de analisar vários desses exemplos, o algoritmo pode traçar um padrão e inferir uma “regra”²⁷ geral, como, por exemplo, os *e-mails* com a frase “Ganhe dinheiro extra” tendem a ser estatisticamente mais prováveis de serem *spam* do que outras mensagens. Com isso, ele pode então usar tais indícios aprendidos para fazer avaliações automatizadas sobre a probabilidade de um novo *e-mail* recebido ser ou não *spam*.²⁸

Em geral, os algoritmos de *machine learning* são capazes de construir automaticamente esses resultados ao inferir as informações por meio da detecção de padrões nos dados. Se esses processos heurísticos estiverem corretos, eles permitirão que os algoritmos façam previsões ou decisões automatizadas envolvendo dados futuros.²⁹ No exemplo citado, o algoritmo detectou um padrão nos dados fornecidos (ou seja, o conjunto de exemplos de *e-mails* de *spam*). Muitos dos *e-mails* que foram sinalizados como *spam* continham a frase: “Ganhe dinheiro extra”. Assim, a partir desse padrão, ele consegue deduzir que *e-mails* com o texto “Ganhe dinheiro extra” tinham maior probabilidade de ser *spam*. Essa generalização pode, portanto, ser aplicada daqui para frente para categorizar, automaticamente, novos

²⁶ Em muitos casos, os algoritmos de *machine learning* são treinados através de conjuntos de dados validados, os quais foram cuidadosamente triados e categorizados por pessoas. Ver, e.g., David Barber, “Bayesian reasoning and machine learning”, 290–96 (2011).

²⁷ O termo “regra” é usado aproximadamente no sentido de “*rule of thumb*” (que significa um princípio ou critério bastante utilizado e que advém das experiências). Isso é importante, porque o *machine learning* é uma técnica indutiva e não dedutiva. Em uma abordagem dedutiva, as regras lógicas gerais (afirmações), que caracterizam o estado do mundo, são expressamente articuladas, e a informação é extraída pela combinação de afirmações de acordo com operações lógicas. Em contraste, em uma abordagem indutiva, modelos do mundo são desenvolvidos ao observar o passado e expressar o estado do mundo (muitas vezes) em probabilidades induzidas pela observação, em vez de como regras gerais. Ver: Daniel Martin Katz, Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry, p. 946.

²⁸ Para ser claro, isto é uma extrema simplificação do *machine learning* para fins ilustrativos. Além dessas, há muitas outras estratégias diferentes. Ver: Mehryar Mohri Et Al., “Foundations of machine learning”, (2012).

²⁹ Toby Segaran, “Programming collective intelligence: building smart web 2.0 applications 3”, (2007).

e-mails recebidos, contendo tal frase, como *spam*. Dessa maneira, o algoritmo tentará detectar outros padrões semelhantes que são comuns entre tais *e-mails* que podem ser usados como uma forma para distingui-los dos outros.

É importante ressaltar que os algoritmos de *machine learning* são projetados para melhorar o seu desempenho ao longo do tempo à medida que recebem mais dados, com o objetivo de construir um modelo interno que consiga solucionar algum fenômeno complexo – como os *e-mails* de *spam* – que, em última análise, permite que o computador tome decisões de classificação precisas e automatizadas. No exemplo, o modelo interno incluiria, cada vez mais, outros exemplos práticos sobre as prováveis características de novos tipos de *spam* que foram recebidos ao longo do tempo – para além do “Ganhe dinheiro extra” –, permitindo que o computador siga classificando novos *e-mails*.

A título de exemplificação, tal algoritmo poderia inferir, a partir de novos exemplos de *spam*, que *e-mails* originários da Bielorrússia³⁰ tendem a ter maior probabilidade de ser *spam* do que *e-mails* recebidos de outros países. Da mesma forma, o algoritmo pode aprender que os *e-mails* enviados por partes com as quais o leitor-usuário já se correspondeu têm menos probabilidade de ser *spam* do que aqueles de remetentes estranhos. Essas descobertas que o algoritmo aprende com a análise de dados adicionais permitem que ele tome melhores decisões automatizadas sobre o que é ou não *spam*.

Conforme ilustrado, os conjuntos de regras³¹ que formam o modelo interno são inferidos a partir do exame e da detecção de padrões dos dados. Por essa razão, tais conjuntos tendem a ser construídos cumulativamente ao longo do tempo, à medida que mais dados são adicionados. Os algoritmos de *machine learning* geralmente desenvolvem métodos heurísticos paulatinamente, examinando cada novo exemplo e comparando-o com os anteriores para identificar semelhanças genéricas que possam ser observadas de forma mais ampla. Por exemplo, um algoritmo pode ter que analisar milhares de exemplos de *spam* antes de detectar um padrão sólido, como perceber que o texto “Ganhe dinheiro extra” é um indício estatístico de provável *spam*.

³⁰ Paul Ducklin, “Dirty Dozen Spam Sending Nations”, NAKED SECURITY (Oct. 17, 2013), <http://nakedsecurity.sophos.com/2013/10/17/dirty-dozen-spam-sending-nations-find-where-you-finished-in-our-q3-spampionship-chart/>.

³¹ É importante notar que esses conjuntos de regras são, muitas vezes, funções matemáticas ou alguma outra estrutura de dados representando o objeto a ser modelado, em vez de uma série de regras formais e gerais. Ver Parag Kulkarni, “Reinforcement and systemic machine learning for decision making...”, p. 2–10.

Por esse motivo, um algoritmo de *machine learning* pode ter um desempenho ruim no início quando possuir apenas alguns poucos exemplos de um determinado fenômeno (como no caso dos *e-mails* de *spam*) para detectar padrões relevantes. Em um momento bem incipiente, seu conjunto de regras internas provavelmente estará bastante subdesenvolvido. No entanto, a capacidade de detectar padrões úteis tende a melhorar à medida que o algoritmo é capaz de examinar mais exemplos do fenômeno em questão. Frequentemente, um algoritmo precisará de dados com muitas centenas ou milhares de exemplos do fenômeno para produzir um modelo interno útil (ou seja, um conjunto robusto de regras preditivas de computador).³²

O exemplo anterior ilustra o que significa “aprender” no contexto do *machine learning*: é a capacidade de melhorar o seu desempenho por meio da detecção de padrões novos ou melhores, a partir da adição de novos dados. Assim, um algoritmo de *machine learning* pode se tornar mais preciso para uma tarefa (como classificar *e-mails* como *spam*) ao longo do tempo pelo fato de que sua configuração permite-lhe aperfeiçoar continuamente seu modelo interno, com a análise de mais exemplos e com a dedução de novos padrões úteis a partir da adição de novos dados.

Logo, é essa capacidade de melhorar o desempenho ao longo do tempo, a partir da análise contínua de dados para detectar padrões úteis, que consiste no principal atributo que caracteriza os algoritmos do *machine learning*. Dessa forma, com base nesse modelo produzido de forma incremental, é que se pode afirmar que esse algoritmo, no seu bom desempenho, pode executar automaticamente uma tarefa - como classificar *e-mails* recebidos como *spam* ou não - com um alto grau de precisão que se aproxima das classificações que um revisor humano em situação semelhante poderia fazer.³³

2.1.2 Detecção de padrões para modelar fenômenos complexos

Há alguns pontos a serem enfatizados sobre o exemplo acima. Primeiro, o *machine learning* geralmente (mas não exclusivamente) desenvolve a sua aprendizagem a partir de um conjunto de exemplos verificados de um determinado fenômeno. Assim, no exemplo anterior, o algoritmo recebeu explicitamente uma série de *e-mails* a qual um ser humano predeterminou como *spam*, possibilitando-o

³²Christopher D. Manning, “Introduction to information retrieval”, 335 (2008).

³³William S. Yerazunis, “The spam-filtering accuracy plateau at 99.9 percent accuracy and how to get past it.”, (Dec. 2004), available at <http://www.merl.com/reports/docs/TR2004-091.pdf> (noting that many spam filters have achieved accuracy rates at over 99.9%).

aprender as suas características através da análise dos exemplos fornecidos. Essa abordagem é conhecida como aprendizagem “supervisionada”, e os exemplos fornecidos, em relação aos quais o algoritmo está sendo treinado para reconhecer padrões, são conhecidos como “conjunto de treino”³⁴. O objetivo desse treinamento é permitir que o algoritmo crie um modelo de cálculo interno, a partir de um determinado fenômeno, de modo que se possa generalizar o aprendizado para se aplicar a novos exemplos nunca vistos desse fenômeno.

Em segundo lugar, esses algoritmos de *machine learning* são capazes de desenvolver, automaticamente, modelos precisos de algum fenômeno – aqui as características do *e-mail* de *spam* –, sem serem explicitamente programados.³⁵ A maioria dos *softwares* são desenvolvidos manualmente, de uma maneira em que os programadores consigam especificar claramente uma série de regras para o computador no intuito de fazê-lo seguir um determinado comportamento desejado. Por exemplo, ao projetar um filtro de *spam* por esse método manual, um programador pode primeiro considerar características que acredita serem típicas de *spam* e, em seguida, programar um modelo com uma série de regras correspondentes a fazerem distinções automatizadas.

Contudo, muitos fenômenos são tão complexos e dinâmicos que são difíceis modelá-los manualmente.³⁶ O problema em uma abordagem manual *bottom-up* para modelar fenômenos complexos e mutáveis (como o *spam*) é a dificuldade de especificar um conjunto de regras *ex-ante* que seja robusto e preciso o suficiente para direcionar um computador a tomar decisões úteis e automatizadas. Por exemplo, um programador pode não pensar em incluir uma regra de que um *e-mail* com origem na Bielorrússia deve ser considerado mais provável como *spam*. Isso demonstra a dificuldade em programar, expressamente, um conjunto de regras para o computador no intuito de produzir automação útil quando se lida com fenômenos complexos e em constante mudança.

Os algoritmos de *machine learning*, por outro lado, são capazes de construir modelos complexos ao detectar, de forma automática, padrões à medida que os dados chegam. Esses algoritmos são poderosos porque, de certa forma, eles se autoprogramam ao longo do tempo, de modo a desenvolver regras para

³⁴Peter Flach, “Machine Learning...”, p. 2.

³⁵Pedro Domingos, “A Few Useful Things to Know About Machine Learning”, COMM. ACM, Oct. 2012, p. 80.

³⁶Pedro Domingos, “A Few Useful Things to Know...”, p. 80.

realizar uma tarefa, em vez de serem programados manualmente com uma série de regras predeterminadas.³⁷ Dessa forma, as regras inferidas a partir de dados analisados possibilitam que o modelo se (re)construa à medida que novos dados são verificados. Utilizando-se do exemplo acima, quando o algoritmo encontra novos exemplos de *spam* em diferentes formas, ele adiciona ao seu modelo interno marcadores adicionais de *spam* com capacidade de detectá-los (como os *e-mails* originários da Bielorrússia). Esse processo incremental, adaptativo e iterativo permite a criação de modelos variados ante fenômenos complexos que, de certa maneira, podem ser muito difíceis para os programadores os predefinirem manualmente.³⁸

Em terceiro lugar, o que tornou o algoritmo de filtragem de *spam* em um algoritmo de *machine learning* foi a sua capacidade de melhorar sua precisão à medida que recebia mais exemplos para analisar. Nesse sentido, estamos usando um significado funcional de “aprendizagem”. Em vez de os algoritmos estarem aprendendo no sentido cognitivo tipicamente associado ao aprendizado humano, eles aprendem a potencializar o seu comportamento no intuito de terem um desempenho melhor à medida que recebem mais dados.³⁹ Assim, no exemplo acima, o algoritmo se tornou mais preciso na identificação de *spam*, pois, ao receber mais exemplos, refinou seu conjunto de regras internas. Dessa forma, podemos conceituar essa mudança como “aprendizagem”, em uma perspectiva funcional, de forma análoga àquela quando associamos o aprendizado humano com a melhoria do desempenho em alguma tarefa.

Em quarto lugar, nota-se que o algoritmo de filtragem acima usou técnicas estatísticas para classificar um *e-mail* como *spam*, de modo a denotar que os algoritmos de *machine learning* são geralmente (embora não exclusivamente) de natureza estatística. Assim, em certo sentido, o *machine learning* não é muito diferente das inúmeras técnicas estatísticas já amplamente utilizadas nos estudos empíricos do direito.⁴⁰ Contudo, uma distinção importante é que muitas das abordagens estatísticas existentes envolvam modelos estatísticos fixos ou de mudança lenta, enquanto

³⁷ Tom Mitchell, “The discipline of machine learning”, Report No. MI-06-Cmu-108 § 1 (2006), disponível em <http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf> (“O *machine learning* se concentra em . . . como fazer com que os computadores se programem (a partir da experiência, além de alguma estrutura inicial).”).

³⁸ Tom Mitchell, “The discipline of machine learning...”. (“A precisão do reconhecimento de voz é maior se se treinar o sistema, do que se se tentar programá-lo à mão.”).

³⁹ I. H. Witten, “Data mining: practical machine learning tools and techniques”, 8 (2d ed. 2005).

⁴⁰ Ver, e.g., David L. Schwartz, “Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases”, 107 Mich. L. Rev. 223 (2008).

o *machine learning* é baseado em algoritmos de computador expressamente projetados para serem dinâmicos e capazes de mudar e se adaptar a cada nova e diferente circunstância à medida que o ambiente de dados muda.

3. Resultados inteligentes sem a inteligência

3.1 *Proxies* e heurísticas para inteligência

O exemplo anterior destinava-se a ilustrar um ponto mais amplo: em alguns cenários é possível realizar tarefas associadas à inteligência humana com algoritmos de computador não inteligentes. Existem certas tarefas, evidentemente, que exigem a inteligência humana porque envolvem habilidades cognitivas de ordem superior, como raciocínio, compreensão, metacognição ou percepção contextual de conceitos abstratos. No entanto, a pesquisa mostrou que algumas dessas tarefas podem ser automatizadas – até certo ponto – por meio do uso de técnicas computacionais não cognitivas que empregam heurísticas ou *proxies* (por exemplo, correlações estatísticas) para a produção útil de resultados “inteligentes”. Por *proxy* ou heurística, refiro-me a algo que pode realizar as funções de outrem ou d’outro sistema eficazmente, substituindo-o, satisfatoriamente, em algum conceito, recurso ou fenômeno subjacente.

Em outras palavras, algoritmos de computador não cognitivos às vezes podem produzir resultados “inteligentes” em tarefas complexas, mesmo sem atingir o nível de cognição humano. Para empregar uma visão funcional da inteligência, tais resultados automatizados podem ser considerados “inteligentes” na medida em que se aproximam daqueles que teriam sido produzidos por uma pessoa em situação semelhante. Essa é uma visão de inteligência orientada para resultados, a partir de uma avaliação baseada em se os resultados produzidos eram sensatos e úteis, e não em se o processo que os produziram foi por meio “cognitivo” *in nature*.

O exemplo de filtragem de *spam* por intermédio de *machine learning* ilustrou essa ideia. Facilmente, podemos auferir que a identificação de um *e-mail* como *spam* por uma pessoa pode envolver uma série de processos cognitivos avançados, pois um usuário humano, para realizar essa tarefa, deve: processar visualmente o *e-mail*, ler, absorver e entender o idioma do texto, contextualizar o significado

do seu conteúdo, raciocinar se o *e-mail* é ou não útil, para, assim, com base nessas avaliações, determinar se o *e-mail* constitui um *spam*.⁴¹

Pode-se, então, concluir que, como a determinação de *spam* envolveu inteligência quando realizada por humanos, a tarefa é inerentemente cognitiva, de modo que, em termos de automação, a maioria dos processos cognitivos complexos descritos acima não são artificialmente replicados por sistemas de computador, em nenhum grau significativo.⁴² Logo, em razão de a identificação de *e-mails* de *spam* envolver cognição e de os computadores não serem capazes de replicar processos cognitivos avançados em nível humano – como, no caso, entender texto arbitrariamente escrito por uma pessoa alfabetizada –, pode-se presumir que não seria possível automatizar um tarefa tão abstrata quanto a de identificar *e-mails* de *spam*.⁴³

Contudo, conforme o exemplo descrito anteriormente, o algoritmo de *machine learning* foi capaz de automatizar a tarefa de filtragem de *spam* por meio de processos não cognitivos, pois, ao utilizar-se do modelo de detecção de padrões, ele conseguiu inferir marcadores de *proxy* eficazes em constatar *e-mails* de *spam*: aqueles com o texto “Ganhe dinheiro extra” ou enviados da Bielorrússia eram, estatisticamente, mais propensos a serem *spam*. Dessa forma, o algoritmo foi capaz de fazer classificações automatizadas úteis e “inteligentes”, aproximando-se do que um usuário humano teria feito depois de ler e compreender o *e-mail*.

No entanto, notavelmente, o algoritmo não se debruçou sobre o significado ou o conteúdo do texto do *e-mail* como faria uma pessoa em situação semelhante, pois, em verdade, nem precisaria,⁴⁴ isto é, o algoritmo não necessita entender conceitos abstratos como “*e-mail*”, “ganhar dinheiro”, “Bielorrússia” ou “*spam*” – da

⁴¹ Ver, e.g., Argye E. Hillis & Alfonso Caramazza, “The Reading Process and Its Disorders”, in *COGNITIVE NEUROPSYCHOLOGY IN CLINICAL PRACTICE* 229, 229–30 (David Ira Margolin ed., 1992) (“Um processo cognitivo, como a leitura, envolve uma série de transformações das representações mentais... Nesta visão, mesmo tarefas cognitivas muito simples, envolverão vários mecanismos de processamento...”)

⁴² Stuart Russell & Peter Norvig, “Artificial intelligence:...”, p. 3–10.

⁴³ Para explicações detalhadas dos limites do Processamento de Linguagem Natural (PNL), a partir da redação deste artigo, ver Stuart Russell & Peter Norvig, “Artificial intelligence:...”, p. 860–67; Robert Dale, “Classical Approaches to Natural Language Processing”, in *HANDBOOK OF NATURAL LANGUAGE PROCESSING* 1, 1–7 (Nitin Indurkha & Frederick J. Damerou eds., 2d ed. 2010); Richard Socher *et al.*, “Semantic Compositionality through Recursive Matrix-Vector Spaces”, in *CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING* § 1 (2012) (notem-se que abordagens particulares de PNL são limitadas e “não captam... a importante qualidade da linguagem natural que permite aos falantes determinarem o significado de uma expressão mais longa com base no significado de suas palavras e nas regras usadas para combiná-las”).

⁴⁴ Toby Segaran, “Programming collective intelligence:...”, p. 4.

maneira como uma pessoa deve – para realizar as classificações automáticas de *spam* de forma precisa. Em vez disso, foi capaz de detectar *proxies* estatísticos em *e-mails* de *spam* e, com isso, produzir resultados úteis e precisos, sem se envolver nos significados ou nos conteúdos subjacentes das palavras constituintes do *e-mail*.

Assim, esse exemplo do filtro de *spam* de *machine learning* ilustrou um ponto bastante profundo: muitas vezes é possível realizar uma tarefa tipicamente associada à cognição sem a necessidade de realizar simulações artificiais de processos intelectuais humanos, mas por meio de algoritmos que empregam heurísticas e *proxies* – como as correlações estatísticas aprendidas por meio de análises de padrões de dados –, que, em última caso, chegam aos mesmos resultados ou a resultados próximos aos que seriam produzidos por uma pessoa inteligente em situação semelhante, através de processos e treinamentos cognitivos avançados.

3.1.1 Assemelhando-se à inteligência por *proxy*

De modo mais geral, o exemplo é ilustrativo de uma estratégia mais ampla que provou ser bem-sucedida na automatização de várias tarefas complexas: detectar *proxies*, padrões ou heurísticas que produzem resultados úteis e de forma confiável em tarefas complexas que, em humanos, normalmente requerem inteligência.⁴⁵ Para um determinado subconjunto de tarefas, pode ser possível detectar *proxies* ou heurísticas que rastreiam de perto o fenômeno subjacente, sem realmente se envolver em toda a sua gama de abstrações, como na maneira como o algoritmo de *machine learning* foi capaz de identificar *e-mails* de *spam* sem a necessidade de entender totalmente o conteúdo e o contexto do texto. Como será discutido na Parte III, esse é o princípio que pode permitir a automatização de certas tarefas abstratas dentro do direito que, quando realizadas por advogados, exigem cognição de ordem superior.

É importante enfatizar, entretanto, que tal abordagem baseada em *proxy* pode ter limitações significativas. Primeiro, essa estratégia pode ser apropriada apenas para certas tarefas para as quais as aproximações são adequadas, sem colocar em risco uma atividade principal, pois muitos problemas complexos – particularmente aqueles que os advogados enfrentam rotineiramente – podem não ser passíveis de tal técnica baseada em heurística. Por exemplo, quando um advogado assiste um cliente corporativo sobre uma possível fusão, que consiste em uma tarefa de tal escala, complexidade e nuances, com tantas considerações, que uma simples abordagem de *proxy* poderia ser inadequada.

⁴⁵Toby Segaran, “Programming collective intelligence...”, p. 1–3.

Em segundo lugar, uma estratégia baseada em *proxy* geralmente pode ter limitações de precisão significativas, em razão de que, como os *proxies* são substitutos de algum outro fenômeno subjacente, eles necessariamente são sub e super abrangidos em relação a uma tarefa a ser desenvolvida perante o fenômeno que estão representando, de modo a, inevitavelmente, poderem produzir falsos positivos e negativos. Ao empregar *proxies* para analisar ou classificar um texto com significado essencial para uma tarefa abstrata, por exemplo, tais algoritmos podem produzir mais falsos positivos ou negativos do que uma pessoa em situação semelhante ao empregar processos cognitivos, conhecimento de domínio e experiência. Assim, permanecendo no exemplo, os filtros de *spam* automatizados podem fazer um trabalho razoavelmente preciso de classificação, mas muitas vezes cometem erros em casos bastante complexos que, contudo, seriam triviais para uma pessoa detectar.⁴⁶ No entanto, sendo as limitações devidamente compreendidas⁴⁷, para certos propósitos (v.g., classificação de *e-mails*), nos quais a eficiência da automação é mais importante do que a precisão, os resultados aproximados podem ser suficientes e úteis.

3.1.2 Os desenvolvimentos nas pesquisas de Inteligência Artificial

A estratégia que acabamos de descrever faz um paralelo com as mudanças na pesquisa sobre inteligência artificial dentro da ciência da computação nas últimas décadas. Na era inicial da pesquisa de IA – dos anos 1950 aos anos 1980 –, muitos pesquisadores se concentravam em tentar replicar os processos cognitivos humanos nos programas de computador. Isso porque acreditavam que, como os humanos utilizam vários processos neurais avançados para enfrentar problemas complexos e abstratos, a maneira de fazer com que os computadores possuíssem inteligência artificial seria criar versões artificiais das funcionalidades do cérebro.⁴⁸

No entanto, mais recentemente, os pesquisadores obtiveram sucesso na automação de tarefas complexas, concentrando-se não na inteligência dos próprios processos automatizados, mas nos resultados que produzem.⁴⁹ Sob essa visão alternativa, se um sistema de computador é capaz de produzir resultados que as pessoas considerariam precisos, apropriados, úteis e vantajosos, poderiam considerá-los “inteligentes” – mesmo que não tenha sido exitoso por meio de versões artificiais dos processos cognitivos humanos.

⁴⁶ William S. Yerazunis, “The spam-filtering accuracy...”, p. 1–5.

⁴⁷ Ver, e.g., Stuart Russell & Peter Norvig, “Artificial intelligence:...”, p. 3–10.

⁴⁸ Stuart Russell & Peter Norvig, “Artificial intelligence:...”, p. 3–10

⁴⁹ Harry Surden, “Computable Contracts”, p. 685–86.

Em geral, essa tem sido a abordagem seguida por muitos sistemas de sucesso de IA dos últimos anos. Esses sistemas usaram *machine learning* e outras técnicas para desenvolver combinações de modelos estatísticos, heurísticas e sensoriais, que não são considerados de natureza cognitiva (na medida em que não replicam a cognição humana), mas que produzem resultados que são úteis e precisos o suficiente para cumprir uma determinada tarefa.⁵⁰ Como já dito, essas abordagens baseadas em *proxy* podem carecer de precisão ou têm outras limitações em comparação com os humanos, para o cumprimento de certas tarefas complexas ou abstratas. No entanto, o principal *insight* é que, para muitas tarefas, as abordagens algorítmicas, como *machine learning*, podem produzir sistemas úteis e automatizados que são “bons o suficiente” para tarefas específicas.

Um bom exemplo desse princípio vem da tarefa de traduzir idiomas. Por muitos anos, a tradução de línguas estrangeiras foi considerada uma tarefa profundamente ligada aos processos cognitivos humanos de ordem superior.⁵¹ Os tradutores humanos fruem de um profundo conhecimento das línguas e de uma grande capacidade de compreensão de conceitos abstratos para traduzir documentos em línguas estrangeiras. Muitos dos primeiros projetos de IA buscavam replicar em computadores várias regras de linguagem que se acredita residirem no cérebro humano,⁵² contudo, utilizando-se de regras *bottom-up*, produziram resultados ruins nas traduções.⁵³

De maneira diversa, as pesquisas mais recentes adotaram uma abordagem diferente a partir do uso de *machine learning* estatístico e do acesso a grandes quantidades de dados, o qual resultou em traduções surpreendentemente boas, sem a necessidade de qualquer tentativa de replicar os processos linguísticos humanos.⁵⁴ O “Google Tradutor”, por exemplo, funciona, em parte, com base nos enormes *corpus* de documentos que os especialistas já traduziram de um idioma para outro, como o caso da Organização das Nações Unidas (ONU), que, ao longo dos anos, empregou tradutores profissionais para traduzir cuidadosamente milhões de seus documentos em vários idiomas, de modo a torná-los disponíveis livremente em

⁵⁰ Stuart Russell & Peter Norvig, “Artificial intelligence:...”, p. 3–10.

⁵¹ Ekaterina Ovchinnikova, “Integration of world knowledge for natural language understanding”, 215–20 (2012).

⁵² Mathias Winther Madsen, “The Limits of Machine Translation”, 5–15 (Dec. 23, 2009) (unpublished Master thesis, University of Copenhagen), disponível em <http://www.math.ku.dk/~m01mwm/The%20Limits%20of%20Machine%20Translation%20%28Dec.%202023,%202009%29.pdf>.

⁵³ Mathias Winther Madsen, “The Limits of Machine Translation...”

⁵⁴ ENCYCLOPEDIA OF MACHINE LEARNING 912–13 (Claude Sammut & Geoffrey I. Webb eds., 2011).

formato eletrônico,⁵⁵ o que possibilitou que os desenvolvedores os aproveitassem para melhorar a tradução automática.

Dessa forma, utilizando-se de correlações estatísticas e de um enorme corpo de dados cuidadosamente traduzidos, os algoritmos automatizados são capazes de criar modelos estatísticos sofisticados sobre o provável significado das frases, produzindo, assim, traduções automáticas de boa qualidade.⁵⁶ Contudo, é importante frisar que os algoritmos que realizam as traduções automáticas não têm nenhuma concepção profunda das palavras que estão traduzindo, nem são programados para entender o significado e o contexto da língua, como um tradutor humano. Em vez disso, esses algoritmos são capazes de usar *proxies* estatísticos extraídos de uma grande quantidade de documentos previamente traduzidos para produzir traduções úteis, sem realmente se envolverem profundamente no idioma.

Embora essa tradução automática muitas vezes fique aquém daquelas feitas por pessoas experts, em termos de precisão e nuances em variados contextos, não sendo suficiente para tarefas que exigem altos graus de precisão (v. g., tradução de contratos jurídicos), o ponto interessante é que, para muitos outros propósitos, o nível de precisão alcançado pela tradução automática pode ser perfeitamente suficiente (v. g., obter uma ideia aproximada do conteúdo de uma página de *web* estrangeira).⁵⁷ Assim, essa automação permite traduções próximas ao que seria realizado por humanos capacitados, sendo úteis naqueles contextos nos quais nenhuma tradução estava disponível anteriormente.

Em síntese, o exemplo de tradução ilustra uma grande estratégia que provou ser bem-sucedida na recente automação da IA: a aplicação de análise de *machine learning* a grandes corpos de dados existentes, para extrair padrões sutis, mas úteis, que podem ser utilizados para automatizar certas tarefas complexas. Dessa forma, essa detecção de padrões em grandes quantidades de dados pode ser usada para criar modelos de computador complexos e diferenciados, no intuito de serem aplicados em problemas que antes eram intratáveis em abordagens manuais de programação.

⁵⁵ Ver: Find Out How Our Translations Are Created, GOOGLE, <http://translate.google.com/about> (último acesso: Feb. 24, 2014).

⁵⁶ Find Out How Our Translations Are Created

⁵⁷ Mathias Winther Madsen, "The Limits of Machine Translation,,," p. 10 (citando Google Translate FAQ, GOOGLE, http://www.google.com/-intl/en/help/faq_translation.html (last visited Mar. 25, 2009)).

4. *Machine learning* e direito

4.1 *Machine learning* aplicado ao direito

Como o *machine learning* tem sido utilizado com sucesso em várias áreas complexas, antes consideradas exclusivas do domínio da inteligência humana, surge a seguinte questão: até que ponto essas técnicas podem ser aplicadas na prática do direito?⁵⁸ Vimos que os algoritmos de *machine learning* geralmente são capazes de construir modelos computadorizados úteis para fenômenos complexos, ao detectar padrões e ao inferir regras a partir de dados. De modo mais geral, também percebemos que as técnicas de *machine learning* são capazes de produzir resultados “inteligentes” em tarefas complexas e abstratas, muitas vezes não por envolvimento direto com o conteúdo conceitual subjacente da informação, mas indiretamente, detectando *proxies* e padrões em dados que levam a resultados úteis. A partir desses princípios, esta última parte do artigo sugere que há um subconjunto de tarefas jurídicas frequentemente executadas manualmente por advogados que, em parte, são potencialmente automatizáveis, com base em técnicas como *machine learning*, devendo-se compreender e contabilizar, contudo, as suas limitações.

Ressalta-se, novamente, que essas tarefas podem ser *parcialmente* automatizáveis, pois o objetivo geral dessas automações não é substituir um advogado, mas sim auxiliá-lo, como, por exemplo, com algoritmos que filtram prováveis dados irrelevantes com intuito de trazer mais eficiência ao labor do jurista. Essa dinâmica será discutida a seguir no caso da automação na revisão de documentos probatórios de litígios, no qual os algoritmos de *machine learning* não são usados para substituir (nem são capazes disso) tarefas cruciais do advogado, como determinar se certos documentos ambíguos são relevantes sob uma determinada normal legal ou se terão significativo valor estratégico em um litígio. Em vez disso, em muitos casos, os algoritmos podem ser capazes de filtrar, de forma confiável, uma grande quantidade de documentos que provavelmente serão irrelevantes, permitindo que o advogado não precise desperdiçar seus recursos cognitivos limitados, analisando-os. Além disso, esses algoritmos podem destacar certos documentos eventualmente relevantes com o escopo de chamar uma maior atenção do advogado. Assim, o algoritmo não substitui o advogado, mas automatiza certos “casos fáceis”, de modo que os seus esforços cognitivos e o seu tempo possam ser

⁵⁸ Isto não quer dizer que outras técnicas de IA não terão impacto sobre a lei. Como escrevi em outro lugar, a IA baseada na lógica está impactando domínios legais, como a contratação. Ver: Harry Surden, “Computable Contracts...”.

conservados para aquelas tarefas que provavelmente exigirão habilidades jurídicas de ordem superior.

Nesse sentido, existem tarefas específicas para as quais os algoritmos de *machine learning* são mais adequados do que outras. Logo, identificando os tipos de tarefas que os algoritmos de *machine learning* executam particularmente bem, podemos, assim, conjeturar sobre situações em que esses algoritmos podem impactar na prática jurídica.

4.2 Modelos preditivos

4.2.1 Previsões judiciais

Os algoritmos de *machine learning* têm sido usados com sucesso para gerar modelos preditivos de certos fenômenos, possibilitando, não obstante, que os resultados desses modelos sejam replicados na prática do direito.⁵⁹

A capacidade de fazer previsões inteligentes e úteis sobre possíveis resultados e responsabilizações legais é uma das principais habilidades da advocacia,⁶⁰ de modo que os advogados são, rotineiramente, provocados a fazerem previsões dentro de uma grande diversidade de cenários jurídicos possíveis. Em geral, quando um cliente apresenta ao advogado um problema jurídico que envolve um conjunto complexo de fatos e finalidades,⁶¹ exige-se do profissional uma combinação de visão crítica, experiência e conhecimento da lei para que se consiga prever, de maneira fundamentada e verossímil, as probabilidades dos resultados em relação a um determinado problema jurídico ou concernente a uma questão geral de responsabilidade, muitas vezes em contextos de considerável incerteza jurídica e factual.⁶² Assim, com base nessas previsões – e em alguns outros fatores –, o advogado consegue responder, de forma adequada, ao cliente sobre as medidas cabíveis.

⁵⁹ Stephen Marsland, “Machine learning: an algorithmic perspective”, 103 (2011).

⁶⁰ Ver, e.g., Tanina Rostain, “Ethics Lost: Limitations of Current Approaches to Lawyer Regulation”, 71 S. Cal. L. Rev. 1273, 1281–82 (1998); Brian Z. Tamanaha, “Understanding Legal Realism”, 87 Tex. L. Rev. 731, 749–52 (2009).

⁶¹ Ver, e.g., Paul Brest & Linda Hamilton Krieger, “Problem solving, decision making and professional judgment”, 29–30 (2010).

⁶² Paul Brest & Linda Hamilton Krieger, “Problem solving...”, p. 29–30 (2010).

Essa habilidade de avaliar de maneira geral a probabilidade de resultados jurídicos e os níveis relativos ao risco da responsabilização, em ambientes de considerável incerteza jurídica e factual, é uma das principais funções que agregam valor a um bom advogado. Como dito, os advogados produzem tais estimativas a partir do uso de uma visão crítica profissional, do conhecimento, da experiência, do treinamento e raciocínio adquiridos ao longo dos anos, além de outras habilidades cognitivas e de intuições.⁶³ No entanto, como escreveu Daniel Katz, tal aptidão estará cada vez mais sujeita a automatização a partir das análises computacionais,⁶⁴ pois, segundo o autor, já existem dados conhecidos que, se bem aproveitados, podem prever melhor os resultados em contextos legais.⁶⁵ Nesse cenário, Katz sugere que a combinação da inteligência humana com as análises computacionais se mostrará, provavelmente, bem superior à análise feita sozinha por uma pessoa ante a grande variedade de questões preditivas.⁶⁶

Nesse sentido, esta parte esboçará uma visão geral de como pode ser essa tarefa preditiva baseada no *machine learning*. Em geral, esse método envolveria o uso de algoritmos de *machine learning* para, a partir de padrões percebidos de dados relativos a cenários jurídicos anteriores, prever resultados em cenários jurídicos futuros. Dessa forma, por meio desse processo, um algoritmo pode ser capaz de detectar *proxies* úteis ou indícios e probabilidades de determinados resultados.

Uma técnica relevante a ser aplicada a esse processo é o método de “aprendizagem supervisionada” discutido anteriormente.⁶⁷ Conforme explicado, a aprendizagem supervisionada consiste em inferir informações, a partir da associação entre dados que foram previamente categorizados por humanos.⁶⁸ Contudo, indaga-se: de onde surgiria esse conjunto de dados? Este surgiria através da criação, pelos escritórios de advocacia, de uma base de dados própria, a partir dos casos vistos anteriormente. Assim, a partir desse levantamento e coleção de informações de clientes anteriores, combinado com outras informações relevantes, como decisões judiciais publicadas, os escritórios poderiam usar algoritmos

⁶³ Ver, e.g., Patrick E. Longan, “The Shot Clock Comes to Trial: Time Limits for Federal Civil Trials”, 35 Ariz. L. Rev. 663, 687 (1993) (“Advogados que possuem mais experiência em julgamentos, e a consequente capacidade de prever resultados com maior precisão, podem cobrar valores de honorários mais vantajosos”).

⁶⁴ Daniel Martin Katz, “Quantitative Legal Prediction...”, p. 912.

⁶⁵ Daniel Martin Katz, “Quantitative Legal Prediction...”, p. 912.

⁶⁶ Daniel Martin Katz, “Quantitative Legal Prediction...”, p. 912.

⁶⁷ Daniel Martin Katz, “Quantitative Legal Prediction...”, p. 912.

⁶⁸ Daniel Martin Katz, “Quantitative Legal Prediction...”, p. 912.

de *machine learning* para construir modelos preditivos referentes a determinados tópicos, como, por exemplo, a probabilidade de responsabilização. Logo, se esses modelos preditivos superaram àquelas feitas pelos advogados em alguns pontos percentuais, eles consistiriam, portanto, em uma valiosa ferramenta para a atividade advocatícia. Dessa forma, a partir da análise de vários exemplos de dados anteriores dos clientes, um algoritmo de *machine learning* conseguiria identificar padrões entre diferentes tipos de informações dos casos judiciais para, com isso, realizar um prognóstico específico.

Para melhor ilustrar, imagine que um escritório de advocacia, que representa demandantes em um caso trabalhista, tenha registrado informações importantes de seus clientes anteriores, em situação semelhante, num banco de dados próprio, de modo que há referências sobre a natureza do incidente e da reclamação, bem como o tipo de empresa onde ocorreu o caso. O escritório, ainda, a partir dos parâmetros citados, analisa aspectos diferentes, como o resultado do caso, se foi resolvido, quanto custou a demanda, qual o juiz do processo, as leis envolvidas e se foi a julgamento etc. Esse conjunto de dados de casos anteriores que o escritório armazenou ao longo dos anos, combinado com outras informações, como demais decisões judiciais ou dados de casos adquiridos por fontes diversas, seria o “conjunto de treinamento”. E, semelhante ao exemplo do filtro de *spam*, o algoritmo de *machine learning* poderia ser treinado, a partir dos exemplos anteriores, para aprender as características mais importantes que indicariam prováveis resultados futuros. Assim, com o passar do tempo e depois de examinar exemplos suficientes de casos de clientes anteriores, um algoritmo de *machine learning* poderia, potencialmente, construir um modelo preditivo, de modo a apontar a influência de fatores que possuem a capacidade de prever determinados resultados.

Com o intuito de melhor esclarecer, imaginemos um algoritmo de aprendizagem que, ao se deparar com casos de discriminação no local de trabalho, em que há um epíteto racial expresso por escrito em um *e-mail*, percebe que, nesses fatos, há uma probabilidade de acordo inicial de 98%, contra um patamar de 60% de acordos nas demais demandas trabalhistas. Embora um advogado possa ter uma intuição profissional semelhante quanto ao acordo antecipado nesses casos, essa informação, amparada por um conjunto de dados, demonstra ser um guia bastante útil para a atividade jurídica.

Portanto, de forma mais eficiente, esses algoritmos podem identificar, diante de uma complexa mistura de fatores no banco de dados, resultados específicos que são muito difíceis ou impossíveis para um advogado detectar usando alguns

típicos métodos de análise jurídica. Nesse sentido, por exemplo, um algoritmo, ao analisar um banco de dados, pode revelar que os casos em que há vários *e-mails* hostis enviados a um funcionário a cada três semanas tendem a ter 15% mais probabilidade de resultar em responsabilização, em comparação aos casos em que *e-mails* semelhantes são enviados em uma periodicidade mais longa, de um ano. Essa conclusão, que observa o período, pode ser difícil para um advogado detectar casualmente, mas é facilmente revelada por meio da análise de padrão de dados. Dessa forma, como o algoritmo recebeu mais e mais exemplos do conjunto de treinamento, ele pode, potencialmente, refinar seu modelo interno, encontrando mais padrões úteis que poderiam melhorar a capacidade do advogado de fazer previsões fundamentadas.

Em suma, aqueles escritórios de advocacia ou entidades litigantes que anseiam por um determinado resultado judicial poderiam, a princípio, alavancar dados de cenários anteriores, tanto dos seus clientes quanto outros dados públicos e privados, para construir modelos preditivos de *machine learning*, estipulando prováveis resultados para complementar o serviço judicial prestado. Em verdade, isso seria transformar, em estatística, o que os advogados costumam fazer intuitivamente⁶⁹, tendo em vista que, ao ficarem perante casos semelhantes repetidas vezes, eles desenvolvem uma compreensão interna e intuitiva de prováveis resultados para esses processos, uma vez que levam em consideração determinados fatores preditivos, em razão de os advogados conseguirem amalgamar sua visão crítica, treinamento, raciocínio, análise, intuição e cognição para compreender os fatos e, assim, realizarem previsões estipuladas para seus clientes. Até certo ponto, os algoritmos de *machine learning* podem desempenhar um papel semelhante, mas complementar, baseado nos dados analisados.

4.2.2 Limitações aos modelos preditivos judiciais de *machine learning*

Existem algumas limitações nos modelos preditivos que devem ser observadas. De um modo geral, o objetivo do uso de *machine learning* é analisar dados anteriores para desenvolver regras que sejam generalizáveis para o futuro. Ou seja, a heurística que um algoritmo detecta ao analisar exemplos passados deve ser útil o

⁶⁹ Isto lembra a citação do grande matemático Pierre-Simon Laplace, que disse há várias centenas de anos: “A teoria das probabilidades nada mais é do que senso comum reduzido ao cálculo; ela nos permite apreciar com exatidão aquilo que as mentes exatas sentem com uma espécie de instinto pelo qual, muitas vezes, elas não são capazes de prestar contas”. (H. C. Tijms, “Understanding probability”, 3-4, (3d ed. 2012), citando LaPlace).

suficiente para produzir resultados precisos em cenários futuros nunca vistos. No exemplo citado, o objetivo era analisar os dados de cenários dos clientes anteriores para associar variáveis (*e-mails* hostis) a resultados específicos (aumento da probabilidade de acordo), com intuito de criar um conjunto de heurísticas que sejam eficazmente preditivas para casos futuros com características um pouco diferentes daquelas do conjunto de treinamento. Dessa forma, o modelo de aprendizagem só é, portanto, útil, à medida que a heurística inferida de casos anteriores possa auxiliar na predição dos novos casos.

Existem, no entanto, alguns problemas bem conhecidos com esse tipo de generalização. Primeiro, um modelo só será útil se a classe de casos futuros tiver características em comum com os casos analisados anteriormente no conjunto de treinamento,⁷⁰ pois se aqueles apresentarem fatos únicos ou incomuns em comparação com estes, o modelo preditivo será menos eficiente. Nesse contexto, as técnicas de *machine learning* podem não ser adequadas para o trabalho de previsão, uma vez que, como no exemplo, nem todo escritório de advocacia terá um fluxo de casos suficientemente semelhantes entre si, de modo que os dados de casos anteriores catalogados contenham elementos que serão úteis para prever resultados futuros. O grau de relacionamento entre casos futuros e passados dentro de um conjunto de dados é uma dimensão importante a ser considerada em relação à extensão em que os modelos preditivos de *machine learning* serão úteis, tendo em vista que tais algoritmos exigem uma amostra relativamente grande de exemplos anteriores para que generalizações robustas possam ser inferidas. Logo, se o número de exemplos (*v.g.*, dados de casos anteriores) é muito pequeno, o algoritmo pode não ser capaz de detectar padrões que sejam preditores confiáveis.

Outro problema comum envolve a supergeneralização, a qual é, essencialmente, o mesmo problema conhecido no campo estatístico, como o *overfitting*.⁷¹ A ideia geral dessa problemática consiste na premissa de que é indesejável que um algoritmo de *machine learning* consiga detectar padrões nos dados de treinamento, pois estes são tão bem ajustados às idiosincrasias ou vieses do conjunto de treinamento, que prejudicam a previsão de novos cenários futuros. Por exemplo, voltando ao exemplo do filtro de *spam*, imagine que os *e-mails* usados como

⁷⁰ Existem outros problemas bem conhecidos com a indução. A indução depende da análise de exemplos do passado para generalizar sobre o futuro. No entanto, sob o chamado problema do “Cisne Negro”, podem nunca ter sido vistos de antemão cenários salientes que podem surgir no futuro. Em tal caso, um modelo treinado sobre dados passados pode não ser suficientemente robusto para lidar com cenários futuros raros ou imprevisíveis. Ver, *e.g.*, Nassim Nicholas Taleb, “The black swan: the impact of the highly improbable”, 1–10 (2d ed. 2010).

⁷¹ Ver Stuart Russell & Peter Norvig, “Artificial intelligence...”, p. 705.

conjunto de treinamento sejam sistematicamente tendenciosos de alguma forma: todos foram enviados de um servidor de dados localizado na Bielorrússia. Um algoritmo de *machine learning* pode inferir, incorretamente, a partir desses dados de treinamento tendenciosos, que os *e-mails* de *spam* são originários apenas da Bielorrússia, ignorando os *e-mails* de *spam* de outros países. Essa inferência é precisa, mas apenas com base nos dados de treinamento específicos usados, o que, contudo, conforme aplicada em todo o mundo (generalização), produz resultados imprecisos já que os dados de treinamento não eram representativos de *e-mails* de *spam* de maneira geral.

Da mesma forma, no contexto da previsão no direito, os dados de casos anteriores, sobre os quais um algoritmo de *machine learning* é treinado, podem ser, sistematicamente, tendenciosos, de forma a produzir resultados preditivos imprecisos. A preocupação, em outras palavras, recai sobre a confiabilidade em um algoritmo que está muito sintonizado com as idiosincrasias dos dados de casos anteriores usados para o treinamento de sua função preditiva. Assim, o algoritmo pode ser capaz de detectar padrões e inferir regras a partir de dados do conjunto de treinamento (v.g., ao examinar os casos anteriores de um escritório de advocacia), mas as regras inferidas podem não ser úteis para fins preditivos, se os dados dos quais os padrões foram detectados foram tendenciosos de alguma forma e não refletindo suficientemente a diversidade de casos futuros que provavelmente aparecerão no mundo real.

Uma questão final que vale a pena mencionar envolve a captura de informações em dados. Em geral, os algoritmos de *machine learning* são apenas tão bons quanto os dados que recebem para analisar, já que constroem modelos estatísticos internos com base nos dados fornecidos. No entanto, em muitos casos de tentativa de previsão no direito, pode haver fatores sutis que são altamente relevantes e que os advogados empregam rotineiramente em suas avaliações profissionais, mas que podem ser de difíceis capturas nos dados formais e analisáveis.

Por exemplo, imagine que existe um conselho administrativo que julga processos disciplinares e que, recentemente, tenha passado por mudança no seu quadro de conselheiros. Um advogado experiente, que trabalhou em uma determinada área por muitos anos, pode estar familiarizado com a composição do conselho e os tipos de casos com que esses indivíduos simpatizam. Assim, tal advogado pode fazer uma recomendação quanto a uma ação em curso para o seu cliente, com base em uma compreensão detalhada dos conselheiros e suas inclinações particulares. Essa informação, percebida por um advogado experiente e que é, frequentemente,

usada em um assessoramento jurídico, é difícil capturar, de forma consistente e precisa, em um modelo de dados. Logo, um modelo de dados que não inclua essas informações de difícil captura, apenas aquelas preditivas, irá produzir resultados inferiores aos produzidos por um advogado.

Da mesma forma, existem certas questões jurídicas cujos resultados podem se voltar à análise de abstrações – como a compreensão da política pública de uma lei e a sua aplicação a um conjunto de fatos –, para os quais pode não haver nenhuma *proxy* de dados adequada. Assim, em geral, se existem certos tipos de informações importantes que são difíceis de quantificar nos dados, e cuja avaliação requer uma análise diferenciada, essas considerações importantes podem estar além do alcance das técnicas de previsão do estado atual de *machine learning*.

4.3 Relações ocultas nos dados

Técnicas de *machine learning* também são úteis para descobrir relações ocultas entre dados existentes que, de outra forma, poderiam ser difíceis de detectar, de modo a permitir que os advogados possam usar o *machine learning* para destacar informações úteis desconhecidas que existem dentro das bases de dados atuais, mas que são ocultadas devido à complexidade. Por exemplo, considere um escritório de advocacia que rastreia dados de clientes e resultados em casos de responsabilidade civil ao longo de vários anos. Um algoritmo de *machine learning* pode detectar correlações sutis, mas importantes, que passariam despercebidas por uma análise humana. Agora imagine, por exemplo, que o algoritmo detecta que a probabilidade de um acordo antecipado é significativamente maior quando o demandado, em um caso de danos pessoais, é um hospital, em comparação com outros tipos de réus. Esse é o tipo de assimilação que um algoritmo de *machine learning* pode detectar, e que pode ser relevante para a prática jurídica, mas que pode ser sutil o suficiente para escapar de uma análise de dados.

Em geral, a mineração de dados existentes em escritórios de advocacia pode dar aos advogados novas informações sobre fatores importantes que afetam os resultados (como aquele sobre o hospital ser o demandado), as quais podem, de outra forma, escapar da análise profissional tradicional. Isso, na verdade, representa um desvio normal da avaliação de informações legais, pois os advogados normalmente confiam na sua intuição interna e na sua experiência para determinar os fatores que tendem a ser relevantes para que se atinja determinados resultados em um caso específico. O *machine learning*, como uma técnica – desde

que se destaque nas correlações de *ferreting* –, pode ajudar a complementar as intuições dos advogados e destacar fatores sutis que poderiam escapar de serem notados. A descoberta de tais informações implícitas, combinada com a análise tradicional dos advogados, pode, potencialmente, impactar e melhorar a prestação do serviço advocatício aos clientes.

4.3.1 Decisões judiciais e relacionamentos de dados

Existem algumas outras aplicações potencialmente profundas dos modelos de *machine learning* que podem revelar relações não óbvias, particularmente na análise de decisões judiciais. Uma premissa do sistema de *common law* dos Estados Unidos é que os juízes são, em geral, obrigados a explicar suas decisões ao emitirem seus julgamentos.⁷² Nesses instrumentos, os juízes normalmente explicam por que decidiram daquela forma ao fazer referência à lei, aos fatos, às políticas públicas e outras considerações sobre as quais o resultado foi baseado.⁷³

Implícito nesse sistema de decisões está a seguinte premissa: o juiz chegou a um determinado resultado pelas razões lá expostas, ou seja, as justificativas que um juiz manifesta explicitamente em uma decisão devem corresponder às motivações reais que o levou a chegar àquele resultado. Logicamente, as decisões judiciais não devem ocorrer por razões diferentes daquelas que foram expressamente declaradas e articuladas ao público. Assim, são indesejáveis as decisões judiciais que não refletem as suas reais motivações, em virtude da probabilidade de haver certas motivações que são consideradas impróprias, ilegais ou indecorosas. Por exemplo, decisões baseadas em animosidade racial são ilegais, e resultados judiciais movidos por puro partidarismo sobre o conteúdo julgado podem ser percebidos como impróprios ou inconvenientes. Além disso, é desejável que as razões judiciais declaradas correspondam às razões reais, porque em um sistema de *common law*, os atores sociais (e advogados) confiam nas decisões judiciais e nas suas justificações declaradas, de modo a permitir fazer previsões sobre resultados legais futuros e para compreender e cumprir a lei.

Como algoritmos de *machine learning* podem ser muito eficientes para detectar relações difíceis de observar entre dados, poderiam, portanto, ser capazes de detectar associações ocultas entre certas variáveis de casos e resultados legais

⁷²Jonathan R. Macey, “Promoting Public-Regarding Legislation Through Statutory Interpretation: An Interest Group Model”, 86 Colum. L. Rev. 223, 253–54 (1986).

⁷³Jonathan R. Macey, “Promoting Public-Regarding Legislation...”

específicos, como, por exemplo, o *machine learning* trazer evidências sugerindo que os juízes estavam baseando suas decisões em considerações que não suas razões declaradas. Isto é, dados analisados dinamicamente poderiam questionar se certos resultados legais foram motivados por fatores diferentes daqueles que foram expressos na decisão judicial.

Em uma pesquisa, que bem ilustra o que foi dito, Theodore Ruger, Andrew Martin e colaboradores construíram um modelo estatístico de resultados da Suprema Corte dos Estados Unidos baseado em vários fatores, incluindo a orientação política da decisão recorrida (isto é, liberal ou conservadora) e o circuito de origem do recurso.⁷⁴ Não apenas o modelo estatístico superou vários especialistas em termos de previsão dos resultados da Suprema Corte, mas também destacou relações de dados subjacentes que não eram completamente percebidos anteriormente.⁷⁵

Por exemplo, a Suprema Corte lida com recursos advindos de muitos Circuitos de Apelação diferentes, de modo que muitos especialistas consideravam o circuito de origem (por exemplo, o Nono ou Sexto Circuitos) da decisão prolatada como menos importante que outros fatores (*v. g.*, o conteúdo de direito do caso) em relação a resultados específicos. Entretanto, a análise dos dados mostrou uma forte correlação entre o circuito de origem e o resultado que a maioria dos especialistas esperavam, com base em sua intuição e discernimento.⁷⁶ Embora essa pesquisa não envolvesse algoritmos de *machine learning* em particular, ela utilizava algumas técnicas estatísticas similares às que poderiam ser usadas em uma abordagem de aprendizagem de máquina.

Essa pesquisa ilustra um ponto básico: a análise estatística das decisões poderia trazer à tona correlações que conseguiriam minar suposições básicas do sistema legal. Se, por exemplo, a análise de dados destacar que as decisões estão altamente correlacionadas com fatores não relacionados com as razões nelas escritas, isso diminuiria a sua legitimidade.⁷⁷ Isso também demonstra, de forma geral, que a heurística

⁷⁴ Andrew D. Martin *et al.*, “Competing Approaches to Predicting Supreme Court Decision Making”, 2 *Persp. On Pol.* 761, 761–68 (2004); Ver também: Theodore W. Ruger *et al.*, “The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision-Making”, 104 *Colum. L. Rev.* 1150, 1151–59 (2004).

⁷⁵ Andrew D. Martin *et al.*, “Competing Approaches...”, p. 761–68.

⁷⁶ Andrew D. Martin *et al.*, “Competing Approaches...”, p. 761–68.

⁷⁷ Para ser claro, isto não é para sugerir que a correlação implica uma causalidade. Ela é perfeitamente consistente para que as decisões da Suprema Corte sejam correlacionadas com um fator não do direito estritamente (por exemplo, circuito de origem) e ainda sejam baseadas em tal determinação. Assim, por exemplo, se um tribunal de circuitos

estatística pode ser preditiva e informativa em um domínio tão abstrato e cheio de incertezas, como o direito, mesmo quando os computadores não se envolvem de fato com o conteúdo jurídico subjacente (por exemplo, o significado e objetivos das leis, doutrinas ou políticas públicas), que é o foco principal dos advogados.

4.4 Classificação e agrupamento de documentos

A prática do direito está entrelaçada com a produção, análise e organização de documentos. Estes incluem decisões judiciais, documentos probatórios, contratos, resumos e muitos outros tipos de documentos jurídicos escritos. Fora do direito, os algoritmos de *machine learning* têm se mostrado úteis na organização, agrupamento e análise automática de documentos para uma série de tarefas.⁷⁸ Esta subseção explorará dois métodos de *machine learning* que podem ser relevantes para a análise e organização automática de documentos jurídicos: 1) a classificação de documentos; e 2) o agrupamento de documentos.

4.4.1 Classificação automatizada de documentos

Em uma tarefa de classificação de documentos, o objetivo de um algoritmo de *machine learning* é classificar, automaticamente, um documento em uma categoria específica e predefinida,⁷⁹ baseada, geralmente, no texto e em outras características do documento.⁸⁰

O exemplo anterior de filtragem de *spam* ilustrava essa ideia de classificação automatizada de documentos, de modo que o algoritmo de *machine learning* tentava classificar um determinado documento de *e-mail* recebido em uma das duas categorias: ou como *spam* indesejado ou *e-mail* desejado. Como vimos, o algoritmo foi capaz de fazer tais classificações automáticas com base nos vários indícios de *e-mails* de *spam* que tinha detectado a partir de exemplos anteriores (v. g., aqueles que possuíam a frase “Ganhe dinheiro extra” ou que o país remetente era

estava consistentemente cometendo erros em sua interpretação da lei, um resultado (invertido) poderia estar altamente correlacionado com um circuito específico, mas esse resultado não significaria, necessariamente, que a decisão estava sendo tomada com base na consideração do circuito de origem.

⁷⁸ Toby Segaran, “Programming collective intelligence...”, p. 6–9.

⁷⁹ Ver, e.g., Kevin D. Ashley & Stefanie Brüninghaus, “Automatically Classifying Case Texts and Predicting Outcomes”, 17 *Artificial Intelligence & L.* 125, 125–65 (2009).

⁸⁰ Kevin D. Ashley & Stefanie Brüninghaus, “Automatically Classifying...”

Bielorrússia). Além disso, o algoritmo foi capaz de “aprender” – refinar seu modelo interno de caracterização dos *e-mails* de *spam* ao examinar mais exemplos – e melhorar sua capacidade de classificação à medida que seu modelo interno e o conjunto de regras de *spam* se tornaram mais sofisticados. Assim, consideramos essa tarefa como “classificação” porque um usuário humano, ao examinar um *e-mail*, estaria essencialmente realizando a mesma tarefa de classificação – decidindo se um determinado *e-mail* recebido está ou não na categoria “*spam*”.

Dentro do direito, há inúmeras tarefas similares que podem ser consideradas como uma questão de classificação de documentos. Para estas, os algoritmos de *machine learning* podem ser úteis e, em alguns casos, já foram implantados.

4.4.2 Classificação de documentos de processos contenciosos

Desde cerca de 2002, os documentos associados a ações judiciais normalmente estão contidos em *websites*, acessíveis eletronicamente, como o sistema de registros “PACER”, que coleciona decisões das Cortes Federais.⁸¹ O núcleo dessas decisões inclui o objeto, as múltiplas partes, o resumo, e as ordens e sentenças emitidas pelo tribunal, podendo haver várias centenas de documentos associados a um caso, como os processos mais complexos. No entanto, ocultos dentro de centenas de documentos do processo, pode haver alguns especialmente importantes – como a petição inicial emendada – cujo acesso é fundamental, mas difíceis de localizar manualmente. Os documentos eletrônicos do tribunal podem se tornar muito longos, com várias páginas de acesso, podendo um documento particularmente importante ser localizado, por exemplo, apenas na página 146 de um total de 300. Assim, encontrar um documento importante, dentro de inúmeras páginas, com documentos de menor interesse, passa a ser bastante difícil.

Ante a tarefa de encontrar e organizar os documentos centrais do caso – que consiste numa classificação de documentos –, podemos, analogamente, utilizar o exemplo da filtragem de *spam*, uma vez que um algoritmo de *machine learning* pode ser treinado para aprender as características que indicam que um determinado documento é uma petição inicial e não, digamos, um simples requerimento da parte. Assim, o algoritmo poderia ser treinado para automatizar as

⁸¹ Ver: Administrative Office of the U.S. Courts, 25 Years Later, PACER, Electronic Filing Continue to Change Courts, THE THIRD BRANCH NEWS (Dec. 9, 2013), <http://news.uscourts.gov/25-years-later-pacer-electronic-filing-continue-change-courts>; Amanda Conley *et al.*, “Sustaining Privacy and Open Justice in the Transition to Online Court Records: A Multidisciplinary Inquiry”, 71 Md. L. Rev. 772 (2012).

classificações dos documentos com base em características, como o texto do documento, e outras meta informações, como os comentários descritivos do *clerk* do tribunal. Dessa forma, os principais documentos eletrônicos do tribunal poderiam ser automaticamente identificados, como “petições iniciais”, “requerimentos” ou “decisões”, por algoritmos de *machine learning*, facilitando às partes a localização de documentos importantes.

Alguns projetos, como o de *Stanford Intellectual Property Litigation Clearinghouse*, empregaram técnicas semelhantes à de *machine learning* a fim de automatizar a organização de documentos de casos muito longos e complexos, com o escopo de facilitar o acesso de documentos cruciais do tribunal.⁸² Logo, de forma mais abrangente, os algoritmos de *machine learning* seriam capazes de fornecer uma classificação inteligente de documentos para ajudar na organização geral.

4.4.3 *E-discovery* e classificação de documentos

Da mesma forma, certos aspectos da produção de provas em litígios podem ser considerados como um problema de classificação de documentos. Nessa fase, cada parte é frequentemente confrontada a um volumoso conjunto de documentos, incluindo *e-mails*, memorandos e outros documentos internos que podem ser relevantes para o direito e para os fatos em questão. Uma tarefa crucial é classificar esses documentos a fim de encontrar aqueles poucos que são realmente relevantes para alguma questão em específico. Assim, por exemplo, em um caso envolvendo fraude de títulos do mercado de capital, alguns *e-mails* que demonstram a intenção de defraudar podem ser extremamente cruciais para provar um elemento da lei. O maior problema é que, nos litígios modernos, o número de documentos apresentados durante a fase de produção de prova pode ser enorme, variando entre as dezenas de milhares e os milhões.

Apenas uma fração extremamente pequena desses documentos é provável que seja relevante para uma questão específica ou para o caso. Em certo sentido, a tarefa é semelhante à de encontrar uma agulha (*v. g.*, *e-mail* com uma prova cabal do ilícito) no palheiro (*v. g.*, entre os milhões de documentos probatórios). Esta tarefa pode ser pensada como uma tarefa de classificação, pois o objetivo é catalogar cada um dos documentos em categorias baseadas na relevância, como, por exemplo,

⁸²Stanford IP Litigation Clearinghouse, STAN. L. SCH., <http://www.law.stanford.edu/organizations/programs-and-centers/stanford-ip-litigation-clearinghouse> (último acesso Jan. 27, 2014).

(para simplificar), altamente relevante, possivelmente relevante, provavelmente irrelevante, altamente irrelevante.

Antigamente, grande parte dessas provas era conduzida manualmente por advogados juniores que faziam a leitura de *e-mails* e usavam seu julgamento para classificá-las, assim como outros documentos, como provavelmente relevantes ou não relevantes.⁸³ Em essência, isso se assemelha à tarefa de classificação de *e-mails spam*, mas com a diferença de que estes consistem em uma classificação dicotômica e binária – um *e-mail* é ou não um *spam*. No entanto, a classificação de um determinado documento probatório de litígio, como relevante ou não relevante, muitas vezes existe após um julgamento contínuo. Alguns documentos podem ser um tanto relevantes, outros altamente relevantes e outros nada relevantes. É nesta última categoria que a automação tem se mostrado altamente útil.⁸⁴

Atualmente, certos aspectos da produção de provas de litígios estão sendo automatizados em parte, muitas vezes por algoritmos de *machine learning*. Similar às tarefas de categorização discutidas anteriormente, em alguns casos, os algoritmos podem categorizar os documentos por probabilidade de relevância (muitas vezes referida como “codificação preditiva” ou “*Technology Assisted Review*”), de modo a serem capazes de filtrar documentos que provavelmente são irrelevantes com base em datas ou nas partes envolvidas, como, por exemplo, quando o algoritmo infere que os *e-mails* que antecederam em dois anos o incidente principal do processo são altamente prováveis de serem irrelevantes. No entanto, há limitações ao que essas técnicas automatizadas podem fazer. Como discutido, os algoritmos não são os mais adequados ou destinados a aplicar o discernimento jurídico em áreas matizadas e incertas. Ao contrário, o papel dos algoritmos, muitas vezes, é de filtrar, para menos, o tamanho da pilha de documentos que, em última análise, necessita de revisão por parte de um advogado, pois, uma vez assinalados, muitos dos documentos ainda requerem a atenção de um jurista a fim de conduzir uma análise jurídica quanto à sua relevância.

⁸³ Ver, *e.g.*, John Markoff, “Armies of Expensive Lawyers, Replaced by Cheaper Software”, N. Y. TIMES (Mar. 4, 2011), <http://www.nytimes.com/2011/03/05/science/05legal.html>.

⁸⁴ Ver, *e.g.*, Vincent Syracuse *et al.*, “E-Discovery: Effects of Automated Technologies on Electronic Document Preservation and Review Obligations”, INSIDE COUNSEL (Dec. 18, 2012), <http://m.insidecounsel.com/2012/12/18/e-discovery-effects-of-automated-technologies-on-e>.

4.4.4 Agrupamento de documentos relacionados

Em um exemplo anterior, o algoritmo de *machine learning* descrito foi usado para classificar documentos em categorias bem compreendidas e predefinidas, tais como “petições iniciais, “requerimentos” ou “decisões”. Em alguns casos, entretanto, os documentos podem ter características em comum, mas as características de unidade dos documentos podem ser desconhecidas ou não óbvias. Neste caso, em que há pontos em comum ocultos ou desconhecidos entre itens, como os documentos, uma abordagem de *machine learning* conhecida como “*clustering* [agrupamento]” pode ser útil.⁸⁵

No agrupamento, um algoritmo de *machine learning* tenta reunir, automaticamente, itens que são similares, por alguma razão, com base em alguma característica comum que se tenha detectado. Em outras palavras, o algoritmo tenta detectar, automaticamente, relações ocultas ou não óbvias entre documentos, que de outra forma não seriam facilmente descobertos, e agrupá-los.

Dessa forma, o algoritmo pode ser usado para descobrir que documentos legais, aparentemente não relacionados, possuem uma ligação uns com os outros de uma forma essencial ou útil. Por exemplo, imagine que existem duas decisões judiciais em duas áreas fundamentalmente diferentes do direito: direito de família e direito de registro de patente. Imagine ainda que as duas decisões compartilham alguma sutileza subjacente referente a um longo debate sobre as melhores estratégias de prática em direito administrativo. Tal conexão entre esses dois casos poderia passar despercebida pelos advogados, uma vez que haveria poucas possibilidades de os profissionais de cada área lerem decisões concernentes a outro campo jurídico. Entretanto, um algoritmo de agrupamento pode ser capaz de encontrar, automaticamente, tal associação e agrupar os documentos através dessa relação não óbvia, detectando um padrão entre um grande conjunto de dados – decisões judiciais.

Agora, considere um outro exemplo em que o agrupamento e a reunião automatizados de documentos podem ter usos dentro do direito. No direito de registro de patentes, os auditores e advogados dessa área tentam, com muito esforço, encontrar documentos publicados que descrevem invenções similares a uma outra patente existente.⁸⁶ Nessa seara jurídica, há uma exigência, por exemplo,

⁸⁵ Ver Rui Xu & Don Wunsch, “Clustering”, p. 2–6 (2008).

⁸⁶ Janice M. Mueller, “Patent Law”, p. 30-40 (4th ed. 2012).

de que um instituto de patente é proibido de emitir uma nova marca, se a invenção reivindicada não for nova.⁸⁷ E uma forma pela qual se determina que uma invenção não é inédita é buscando documentos, ou outras provas, de “estado da técnica”, o qual descreve a invenção, mas que antecede o pedido de patente e que consiste, geralmente, em artigos publicados anteriormente em revistas científicas, patentes ou pedidos de patente que indicam que aquela determinada invenção foi criada anteriormente, portanto, não é nova.

No entanto, diante do enorme volume de patentes publicadas e de revistas científicas, é uma tarefa difícil encontrar documentos específicos de “estado da técnica”, em um cenário muito amplo, que provariam que uma invenção foi concebida anteriormente. Contudo, a problemática centra em encontrar esse documento, o qual é necessário para determinar, automaticamente, a relação entre um pedido de patente e o documento anterior do “estado da técnica”. Assim, o agrupamento de documentos de *machine learning* pode, potencialmente, ajudar a tornar a busca de documentos relacionados ao estado da técnica mais automatizada e eficiente, por meio da reunião de documentos relacionados ao pedido de patente em questão. Em verdade, e de modo mais geral, o agrupamento automático de documentos poderia ser útil em outras áreas do direito, nas quais a busca de documentos relevantes, em grandes amontoados, é crucial.

5. Conclusão

O artigo buscou uma abordagem da ciência da computação, conhecida como *machine learning*, e seu potencial impacto sobre a prática jurídica. Consistiu, na verdade, em um a visão geral de que, como a tecnologia atual de IA não pode corresponder à análise abstrata e às habilidades cognitivas avançadas, mostradas rotineiramente por advogados treinados, as técnicas atuais de IA podem ter pouco impacto sobre o direito, exceto por avanços tecnológicos significativos. No entanto, este artigo mostrou que, fora do direito, as técnicas de IA – particularmente a *machine learning* – têm sido aplicadas com sucesso a problemas que, tradicionalmente, se pensava que exigiam cognição humana.

Nesse sentido, sugeriu-se que, da mesma forma que ocorre em outros ramos, há uma série de tarefas dentro da prática jurídica em relação às quais as avaliações

⁸⁷ 35 U.S.C. § 102(a) (2006 & Supp. V 2011).

estatísticas, dentro do âmbito das técnicas de *machine learning*, são susceptíveis de terem impactos, apesar da incapacidade de replicar tecnologicamente a cognição avançada dos advogados. Assim, em percepção geral, conclui-se que as avaliações estatísticas e outras avaliações automatizadas de dados baseadas em heurística podem, às vezes, produzir resultados em tarefas complexas que, embora potencialmente menos precisas que os resultados produzidos pelos processos cognitivos humanos, podem na verdade ser suficientemente precisas para certos propósitos que não exigem níveis extremamente altos de precisão e exatidão.

Referências

Administrative Office of the U.S. Courts, 25 Years Later, PACER, Electronic Filing Continue to Change Courts, THE THIRD BRANCH NEWS (Dec. 9, 2013), <http://news.uscourts.gov/25-years-later-pacer-electronic-filing-continue-change-courts>

Amanda Conley *et al.*, “Sustaining Privacy and Open Justice in the Transition to Online Court Records: A Multidisciplinary Inquiry”, 71 Md. L. Rev. 772 (2012)

Andrew D. Martin *et al.*, “Competing Approaches to Predicting Supreme Court Decision Making”, 2 Persp. On Pol. 761, (2004)

Argye E. Hillis & Alfonso Caramazza, “The Reading Process and Its Disorders”, in COGNITIVE NEUROPSYCHOLOGY IN CLINICAL PRACTICE 229, (David Ira Margolin ed., 1992)

Brian Z. Tamanaha, “Understanding Legal Realism”, 87 Tex. L. Rev. 731, (2009)

Burkhard Bilger, “Auto Correct: Has the Self-Driving Car at Last Arrived?”, NEW YORKER, Nov. 25, 2013

Cass R. Sunstein, Kevin Ashley, Karl Branting & Howard Margolis, “Symposium: Legal Reasoning and Artificial Intelligence: How Computers Think Like Lawyers”, 8 University of Chicago Law School Roundtable 1 (2001)

Christopher D. Manning, “Introduction to information retrieval”, (2008)

Daniel Martin Katz, “Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry”, 62 Emory L.J. 909 (2013)

David Barber, “Bayesian reasoning and machine learning”, (2011)

David Bellos, “Is that a fish in your ear? Translation and the meaning of everything”, (2011)

David E. Sorkin, “Technical and Legal Approaches to Unsolicited Electronic Mail”, 35 U.S.F. L. Rev. 325 (2001)

David L. Schwartz, “Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases”, 107 Mich. L. Rev. 223 (2008)

Ekaterina Ovchinnikova, “Integration of world knowledge for natural language understanding”, (2012)

ENCYCLOPEDIA OF MACHINE LEARNING, (Claude Sammut & Geoffrey I. Webb eds., 2011)

Find Out How Our Translations Are Created, GOOGLE, <http://translate.google.com/about> (ultimo acesso: Feb. 24, 2014)

H. C. Tijms, “Understanding probability”, (3d ed. 2012)

Harry Surden, “Computable Contracts”, 46 U.C. Davis L. Rev. 629 (2012)

I. H. Witten, “Data mining: practical machine learning tools and techniques”, (3d ed. 2011)

I. H. Witten, “Data mining: practical machine learning tools and techniques”, (2d ed. 2005)

Janice M. Mueller, “Patent Law”, (4th ed. 2012)

John Markoff, “Armies of Expensive Lawyers, Replaced by Cheaper Software”, N. Y. TIMES (Mar. 4, 2011), <http://www.nytimes.com/2011/03/05/science/05legal.html>

Jonathan R. Macey, “Promoting Public-Regarding Legislation Through Statutory Interpretation: An Interest Group Model”, 86 Colum. L. Rev. 223, (1986)

Karl Okamoto, “Teaching Transactional Lawyering”, 1 Drexel L. Rev. 69 (2009)

Kevin D. Ashley & Stefanie Brüninghaus, “Automatically Classifying Case Texts and Predicting Outcomes”, 17 Artificial Intelligence & L. 125, 125–65 (2009)

Lawrence Maisel, “Predictive business analytics: forward looking capabilities to improve business performance”, (2014)

Mathias Winther Madsen, “The Limits of Machine Translation”, 5–15 (Dec. 23, 2009) (unpublished Master thesis, University of Copenhagen), disponível em <http://www.math.ku.dk/~m01mwm/The%20Limits%20of%20Machine%20Translation%20%28Dec.%2023,%202009%29.pdf>.

Mehryar Mohri *Et Al.*, “Foundations of machine learning”, (2012)

Nassim Nicholas Taleb, “The black swan: the impact of the highly improbable”, (2d ed. 2010)

Parag Kulkarni, “Reinforcement and systemic machine learning for decision making 1–2”, (2012)

Patrick E. Longan, “The Shot Clock Comes to Trial: Time Limits for Federal Civil Trials”, 35 Ariz. L. Rev. 663, (1993)

Paul Brest & Linda Hamilton Krieger, “Problem solving, decision making and professional judgment”, (2010)

Paul Ducklin, “Dirty Dozen Spam Sending Nations”, NAKED SECURITY (Oct. 17, 2013), <http://nakedsecurity.sophos.com/2013/10/17/dirty-dozen-spam-sending-nations-find-where-you-finished-in-our-q3-spampionship-chart/>.

Pedro Domingos, “A Few Useful Things to Know About Machine Learning”, COMM. ACM, Oct. 2012

Peter Flach, “Machine Learning: The art and science of algorithms that make sense of data 3”, (2012)

Richard Socher *et al.*, “Semantic Compositionality through Recursive Matrix-Vector Spaces”, in CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, (2012)

Robert Dale, “Classical Approaches to Natural Language Processing”, in HANDBOOK OF NATURAL LANGUAGE PROCESSING 1, 1–7 (Nitin Indurkha & Frederick J. Damerau eds., 2d ed. 2010)

Rui Xu & Don Wunsch, “Clustering”, (2008)

Stanford IP Litigation Clearinghouse, STAN. L. SCH., <http://www.law.stanford.edu/organizations/programs-and-centers/stanford-ip-litigation-clearinghouse> (último acesso Jan. 27, 2014)

Stephen Marsland, “Machine learning: an algorithmic perspective”, (2011)

Stuart Russell & Peter Norvig, “Artificial intelligence: a modern approach”, (3d ed. 2010)

Tanina Rostain, “Ethics Lost: Limitations of Current Approaches to Lawyer Regulation”, 71 S. Cal. L. Rev. 1273, (1998)

Theodore W. Ruger *et al.*, “The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision-Making”, 104 Colum. L. Rev. 1150, (2004)

Toby Segaran, “Programming collective intelligence: building smart web 2.0 applications 3”, (2007)

Tom Mitchell, “The discipline of machine learning”, Report No. ML-06-Cmu-108 § 1 (2006), available at <http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf>

Vincent Syracuse *et al.*, “E-Discovery: Effects of Automated Technologies on Electronic Document Preservation and Review Obligations”, INSIDE COUNSEL (Dec. 18, 2012), <http://m.insidecounsel.com/2012/12/18/e-discovery-effects-of-automated-technologies-on-e>.

William S. Yerazunis, “The spam-filtering accuracy plateau at 99.9 percent accuracy and how to get past it”, (Dec. 2004), available at <http://www.merl.com/reports/docs/TR2004-091.pdf> (noting that many spam filters have achieved accuracy rates at over 99.9%).

Legislação citada

35 U.S.C. § 102(a) (2006 & Supp. V 2011).